Exploring an alternative scheme of room acoustics specification
Andor T Fürjes

To cite this version:

HAL Id: hal-03235306
https://hal.science/hal-03235306
Submitted on 26 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
EXPLORING AN ALTERNATIVE SCHEME OF ROOM ACOUSTICS SPECIFICATION

Andor T. Fürjes\textsuperscript{1,2}

\textsuperscript{1} aQrate Acoustics Ltd., H8083 Csákvár, Hungary
furjes.andor@aquire.hu
\textsuperscript{2} Animative Ltd., H1043 Budapest Dugonics street 11., Hungary
andor.f@animative.eu

ABSTRACT

To define a specification scheme for room acoustics is a basic task that usually ends up by relying mainly on mean reverberation time tolerances. Many papers investigate how reverberation time correlates (or not) to other measures. One shall also consider robustness and ease of measurement methods of selected parameters when creating regulations.

While aiming for a practical specification scheme, the author found that instead of (late) reverberation times the early decay time parameter should gain more emphasis in general cases: it correlates well with speech definition metrics and it is more robust to measure or model.

In addition, measurement data suggests that instead of arithmetically averaging results at adjacent frequency bands (mean reverberation time), evaluation of a pre-filtered (e.g. A-weighted) wide band signal would yield a more practical and robust single figure.

The paper presents both measurement and calculated results.

1. INTRODUCTION

Room acoustics standards and legislations have a long history several countries and there are numerous papers that sum up or compare their set of requirements or just evaluate their applications.

As of writing this paper, Hungary has no legislation on room acoustics parameters directly. In everyday practice, engineers of room acoustics have to refer to foreign national standards and sometimes have to convince clients in order to avoid obvious mistakes. Theoretically if there is no explicit instruction of the client, room acoustics treatment can be omitted without the legislative force.

The situation is not tragic, however. In most cases clients or project engineers are usually aware of the importance of room acoustics mostly due to experience with technical specifications of foreign developers or former failures.

In 2016 a private initiative started discussion of a national standard and since then a continuous debate is in process. It would have been easy to just refer to a selected existing foreign standard, but as one said in the film English Patient: ‘You’re Hungarian. You always disagree.’

This paper presents some of the doubts and results of investigations to pursue a meaningful room acoustics specification scheme.

Please note, that the paper looks at general situations only, rather than special or artistic cases (i.e. concert halls, opera houses, theatres).

2. INITIAL THOUGHTS

Basically, any general room acoustics specification shall set simply the minimum requirements in order to avoid fundamental mistakes. At the same time clear instructions shall be given to make clear what is better and what is not acceptable.

Some would say, that instead of minimum requirements, optimum requirements shall be set. However, despite the numerous papers on topics of room acoustics there seems to be no global agreement on what is exactly ‘optimal’. Also, setting requirements to a so-called optimum suggests that anything else is simply bad.

There might be trends towards using a unified letter denoted classification scheme, but this should be treated carefully. First of all, for any engineer it is awkward using letters instead of numbers. Secondly, preferred value ranges and set of parameters still seem to change over time as understanding of acoustic impression evolves. Finally, where a trained ear under laboratory conditions might hear differences we call JND (just noticeable difference), a client would appreciate a more reasonable metric to make decisions, like JMD (just meaningful difference, [1]).

3. REVERB OR DECAY

Most room acoustic schemes are based on or list at first place some sort of reverberation time. Surely, if conditions of the Sabine equations hold, and there are no irregularities in a perfectly exponential decay, then a single reverberation time says almost all one needs to know.

While there are papers supporting reverberation time as a fundamental parameter (e.g. [2] showing good correlations for concert halls), a lot of papers raise doubts and report low correlation of reverberation times to other important parameters (e.g. gain, definition, speech intelligibility, etc., see e.g. [7]) in more general use.
Causes of deviations are usually identified as deviations from the ideal exponential decay for any reason. Either the room is too damped, sizes are disproportionate, absorption is unevenly distributed or surface diffusion is too low.

However, there might be another obvious reason: reverberation time by definition tells only the length of the decay but not its quality. How can one tell the quality of a book by only measuring its thickness?

Since direct measurement of reverberation time is hardly feasible, measurement standards (e.g. [9]) suggest to estimate reverberation time by best fitting lines between pairs of points of the energy decay curve (see Figure 1.), which can be calculated from the impulse response $h(r)$:

$$EDC(t) = 10 \cdot \log_{10} \int_{0}^{t} \frac{h^2(r)dr}{\int_{0}^{\infty} h^2(r)dr}$$  \hspace{1cm} (1)$$

![Figure 1. Schematic drawings of early decay time (blue) compared to (late) reverberation time (red) on the EDC. Shaded areas show range of possible EDCs with the same decay properties.](image)

Energetically speaking this means, that reverberation time estimates characterize only the last and lesser portion of the impulse response (the -5 dB…-35 dB is about 32% of the energy of the impulse response). By comparison, early decay time (or EDT$_{10}$) is the length of the first 10 dB drop in the decay, which characterizes the first 90% energetic portion of the impulse response.

It is known, that impression of room acoustic quality is based by the very first part of the impulse response. If we accept this, why do we expect that late reverberation times would always correlate well with parameters that characterize the early and larger portion of the impulse response? In case of EDT$_{10}$ compared to late measures: what do we expect to tell from the rest of the impulse response, when 90% of the energy is already delivered?

Choosing EDT$_{10}$ as a qualitative measure seems reasonable, because it is relatively easy to evaluate and even less prone to in situ background noise than T$_{10}$, T$_{20}$ or T$_{30}$ (see e.g. [3]).

4. CORRELATIONS

To check usability and correlations of EDT$_{10}$, both measurements and results of computer models were analyzed.

4.1 Measurement results

4.1.1 Reverberation time vs. early decay time

Measurement results of a broad range of situations (600…800 measurements, V = 50…20,000 m$^3$) is shown in scatter plots of Figure 2. In the figure T$_{20}$ is T$_{2.25}$ (see [9]) and “m2” denotes arithmetic average of results in 500 Hz and 1 kHz octave bands. For better resolution, both horizontal and vertical axes are shown with log$_{10}$ scale.

Please note, that most measurements were taken using omnidirectional loudspeakers, but some measurements results were collected using directional loudspeakers and even turned in different directions, which caused even larger fluctuations (see ranges of EDT$_{10}$ at T$_{20,m2}$ 0.8 s and 6.4 s).

Figure 3 shows the same measurements but instead of averaging results at 500 Hz and 1 kHz octave bands, A-weighted wide band impulse responses were evaluated. Here the correlation is more clear and stable.

Overall statistical results are summarized in Table 1.

<table>
<thead>
<tr>
<th>Measurement</th>
<th>EDT$<em>{10,ave}$/T$</em>{20,ave}$</th>
<th>EDT$<em>{10,A}$/T$</em>{20,A}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>for each measurement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>min</td>
<td>0.31</td>
<td>0.29</td>
</tr>
<tr>
<td>max</td>
<td>2.28</td>
<td>1.55</td>
</tr>
<tr>
<td>average</td>
<td>0.95</td>
<td>0.82</td>
</tr>
<tr>
<td>std. dev.</td>
<td>0.28</td>
<td>0.17</td>
</tr>
<tr>
<td>for each room</td>
<td></td>
<td></td>
</tr>
<tr>
<td>min</td>
<td>0.56</td>
<td>0.51</td>
</tr>
<tr>
<td>max</td>
<td>1.91</td>
<td>0.99</td>
</tr>
<tr>
<td>average</td>
<td>1.01</td>
<td>0.87</td>
</tr>
<tr>
<td>std. dev.</td>
<td>0.26</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Table 1. Overall statistic figures of EDT$_{10}$/T$_{20}$ ratios.

Figure 3 shows the same measurements but instead of averaging results at 500 Hz and 1 kHz octave bands, A-weighted wide band impulse responses were evaluated. Here the correlation is more clear and stable.

Overall statistical results are summarized in Table 1.

These observations suggest that evaluation of pre-filtered wide band impulse responses might be more suitable than averaging octave band results, also because evaluation is faster and more robust against background noise levels.

Even if EDT$_{10}$/T$_{20}$ can highly vary from measurement to measurement, taking ratios of arithmetic averages for whole rooms give more stable results. According to results in Table 1, EDT$_{10,A}$/T$_{20,A}$ is expected to be lower than T$_{20,A}$ and more than 85% of cases between 80…100% of T$_{20,A}$.
4.1.2 STI vs. reverberation time and early decay time

Figure 4 and 6 show scatter plots of $T_{20,m2}$ and $EDT_{10,m2}$ vs. STI (EN 60268-16, indirect calculation, no masking, standard spectra), where only the horizontal axis is shown in logarithmic scale.

Figure 5 and 7 show the same but with A-weighted mean reverberation time ($T_{20,A}$) and early decay time ($EDT_{10,A}$).

As one can see, $EDT_{10,A}$ correlates to STI much better, namely regression model of

$$STI_{estimate} \approx -0.38 \cdot log_{10}(EDT_{10,A}) + 0.59 \quad (2)$$

yields $R^2 = 0.93$.

These observations suggest that a room response can degrade intelligibility down to a quite certain point, and that intelligibility can be increased e.g. if the speaker has higher directionality or the speaker is closer to the listener.

In addition, both graphs seem to provide a general ‘worst case estimate’ for STI, true in at least 95% of all measured cases:

$$STI > -0.45 \cdot log_{10}(T_{20,m2}) + 0.57 \quad (3)$$

$$STI > -0.34 \cdot log_{10}(EDT_{10,A}) + 0.54 \quad (4)$$
more meaningfully. Finally, reverberation time by itself seems unable to become a reliable estimator of intelligibility (or clarity).

Figure 6. Comparison of mean early decay times (EDT_{10,m2}) and standard speech transmission index (STI).

Figure 7. Comparison of A-weighted early decay times (EDT_{10,A}) and standard speech transmission index (STI). Estimates are in Eq (2) and (4).

4.1.3 Single number ratings

While a single number rating is comfy for simple specifications, it is not clear, why certain types of mean reverberation times are believed to be better or worse for the purpose.

In addition to the mean “m2” it is possible to add results calculated in 2 kHz (T_{m3} or T_{m4}) or even 125 Hz (T_{m4}) octave bands to get a mean value, but according to measured results, there seems to be no real benefit of using more octave bands.

Figure 8 compares different single number ratings of reverberation time and it is obvious, that only T_A has a very slightly differentiated behavior.

Figure 8. Comparison of single number ratings of reverberation times T_{500Hz}, T_{m2}, T_{m3}, T_{m4} and T_A.

4.1.4 Spectral tolerances

One might argue that including more frequency bands in the mean might provide a better view on spectral balance of reverberation time. In the view of the author however, this statement is misleading, because an averaged spectrum cannot substitute a properly chosen spectral tolerance requirement.

Experience does not support that inclusion of additional frequency bands is able to express evenness of spectral balance, so a spectral tolerance shall be applied independently from the definition of the single number rating.

In general, there are two main reasons to specify some kind of spectral tolerance:

a) to avoid excessive high/low ratios
b) to avoid excessive single band deviation.

Figure 9 shows relative deviations from T_{20,m2} mean results for each measurement considering all measurements. Figure 10 shows the same but for the case of early decay time. Natural tendencies of reverberation and decay times in untreated rooms seem to fit within an acceptable tolerance. High (1kHz, 2kHz, 4kHz) and low (125 Hz, 250 Hz, 500 Hz) mean reverberation time ratios showed an average of 0.87 with a deviation of 0.23.

Usage of spectral tolerances is validly advised for musical applications or if high SPL is considered in the room below 250 Hz bands.

However, in general cases when considering only normal speech signal (see e.g. [4]) and typical (L_{Aeq}=30 dBA) background noise spectra, the perceived signal to noise ratio can be much smaller than 30 dB at 125 Hz and below.
(see also [5]), making the use of late reverberation metrics and <250 Hz frequency tolerances somewhat meaningless. On the other hand, EDT\textsubscript{10} seems still more meaningful at low frequencies, because it expresses level change rates that are surely perceptible in the cases of normal speech dynamics under ordinary conditions.

![Figure 9](image)

**Figure 9.** Deviation of reverberation times at each frequency band relative to $T_{\text{m}2}$ mean values. Percentages denote statistical percentiles.

![Figure 10](image)

**Figure 10.** Deviation of early decay times at each frequency band relative to $T_{\text{m}2}$ mean values. Percentages denote statistical percentiles.

### 4.1.5 Room conditions

Usually room acoustic specifications are meant either for furnished/unfurnished or occupied/unoccupied situations. This is a rather simplified scheme, because room acoustics is significantly different also depending on whether a room is inhabited or uninhabited.

In certain applications (e.g. kindergarten rooms, elementary schools) operational conditions can significantly differ from what a room acoustic engineer can validly consider or measure before the rooms is used.

The above can explain, why unoccupied kindergarten rooms could result $T_{\text{m}2}<0.6$ s and unoccupied elementary school classrooms could result $T_{\text{m}2}<0.7$ s reverberation times without any intentional room acoustic treatment.

Since vocal comfort can degrade by adding too much damping (see [6]), consideration of expected additional absorptions is advised in such cases.

In the view of the author, specifications for partial or fully occupied situations shall be avoided due to the uncertain nature of absorption occupants can present and the lack of controlled environment during verification. Besides, specifying the unoccupied condition is in most cases practically the worst-case scenario specification.

![Table 2](image)

**Table 2.** Differentiation of conditions of a room to measure.

Due to the above considerations, comparable room acoustic specifications shall be formed for furnished, unoccupied, uninhabited situation, without any furniture that can be removed under normal operational conditions (e.g. an event hall can be used without mobile chairs, but a class room without chairs is not functional).

### 4.2 Modelling results

Numerous papers (e.g. [7]) demonstrate how Sabine, Eyring or other statistical reverberation time formula cannot reliably correlate other monaural acoustic parameters like G, clarity, definition or STI, if conditions (e.g. room proportions, distribution of absorption) are not ideal.

To test a wide range of scenarios, a series of modelling calculations were carried out on 2 types of geometries (see Figure 11). Both geometries have 1000 m$^3$ volume, 200 m$^2$ floor area and 5 m height. Boundaries are divided to 5×10 m (50 m$^2$) patches. In case of Geometry #B patches 12 and 13 are the same. Source is at (6.0;6.0;1.5) m coordinates. Omni receivers of 1×1 m raster are at 1.1 m height in two parts of the room, each 8×8 m (1 m away from side walls).

Initially all boundaries are reflective (smooth concrete) and only 2 patches (so 100 m$^2$ in total) are set to absorptive (4” glass fiber). Different combinations of absorptive patches (see Table 3) are calculated at 5%, 10%, 20% and 40% minimum diffusity settings (128 model runs total).
Finally, mean (m2) and A-weighted results of EDT₁₀, T₂₀, G, C₅₀, C₈₀ parameters are compared with each other and STI.

Modelling calculations were carried out using EASE Aura Module (4.4.6) and PETRA¹. Both handle non-diffuse (specular) and diffuse reflections and overall diffuse reflection ratios can be set for each calculation.

Due to the high number of results, only some examples are shown in Figure 12 (test geometry #A) and Figure 13 (test geometry #B). Bars show average of mean values of different models with 5%, 10%, 20% and 40% diffuse settings.

Table 4 shows relation of different parameter results expressed by coefficient of determination (R²) of linear regression.

Reverberation times calculated by the Sabine formula is 1.48/1.47 s for the #A/#B geometries and A…Q cases. Reverberation times according to the Norris-Eyring formula are 1.36/1.36 s for the same.

Results confirmed the following observations:
- T₂₀,mf or T₂₀,A has poor or weak correlation to G or C₅₀ and only modest correlation to STI;
- EDT₁₀,mf or EDT₁₀,A has poor correlation to G, modest correlation to C₅₀ and firm correlation to STI, correlation is better for EDT₁₀,A;
- spatial average minus standard deviation of STI seems to correlate to both T₂₀ and EDT₁₀;
- C₅₀ is highly correlated to STI;
- G is not correlated to decay length or clarity and its spatial average depends mainly on absorption quantity, but barely depends on its position or room shape;
- settings of diffusity affect T₂₀ more than EDT₁₀; and arithmetic mean values more than A-weighted values.

Results also show, that reverberation or decay lengths are lowest, STI and clarity are highest, while G is lowest if absorption is arranged according D, I and K versions. The opposite is true for arrangement M. Other arrangements seem to perform broadly the same.

Generally speaking and not surprisingly, the more evenly absorption is distributed, the more efficiently it reduces reverberation and its consequences.

### 5. DISCUSSION

It is known, that early decay time (EDT₁₀) is subjectively more important than reverberation time (see e.g. [9]) and that maybe reverberation times shall be taken only as quantitative parameter.

This paper presented experience from wide range of measurements and a systematic series of modelled scenarios and concluded, that EDT₁₀ is indeed a better candidate to represent fundamental acoustic qualities of rooms of general use:
- firm correlation to clarity and intelligibility;

---

¹ PETRA is an experimental acoustic modelling environment, here combining only phased beam tracing and radiosity methods.
- measurement is less sensitive to background noise and
does not necessarily require impulse response
measurements (easier in situ evaluation possible).

There are problems to solve and concerns to clear, though:
- there is no simple formula to predict EDT;
- EDT measurement and evaluation practice: correct
direct sound detection and higher spatial variance can
cause ambiguity.

Figure 12. Modelling results summarized and compared
(test geometry #A).

Figure 13. Modelling results summarized and compared
(test geometry #B).
Table 4. Relation of different parameters expressed by coefficients of determination, calculated at 10% diffuse ratio.

<table>
<thead>
<tr>
<th>T20</th>
<th>εRD</th>
<th>G</th>
<th>C50</th>
<th>STI std</th>
</tr>
</thead>
<tbody>
<tr>
<td>mf</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>1.00</td>
<td>0.96</td>
<td>0.99</td>
<td>0.95</td>
<td>0.94</td>
</tr>
<tr>
<td>1.00</td>
<td>0.95</td>
<td>0.99</td>
<td>0.94</td>
<td>0.93</td>
</tr>
</tbody>
</table>

To estimate EDT10 of a room, the rate of level change caused by the effect of the first absorptions is calculated:

\[
\Delta T = \frac{10 \log_{10}(1 - \sigma_{\text{avg}})}{\log_{10} \frac{\sigma_{\text{avg}}}{10}} = \frac{10 \cdot \frac{\sigma_{\text{avg}}}{10}}{\log_{10} \frac{\sigma_{\text{avg}}}{10}} \left( \frac{\text{dB}}{\text{sec}} \right)
\]

where \( \sigma_{\text{avg}} \) is the average absorption coefficient, \( l_{\text{avg}} \) is the mean free path between reflections. For a -60 dB decay:

\[
T_{60dB} \approx -\frac{60\text{dB}}{\Delta T/\Delta t} = -\frac{24\ln(10)}{c} \frac{V}{\left[-5\cdot l_{\text{avg}}(1-\sigma_{\text{avg}})\right]} \left( \text{sec} \right)
\]

which is exactly the Norris-Eyring equation.

Of course, more detailed estimates can consider effects of directional behavior and source-receiver distance (see [2]).

As suspected, a pre-filtered (e.g. A-weighted) wide band evaluation might be more meaningful than an arithmetic average of frequency dependent evaluation results. Oddly however, simple averages can proximate well:

\[
T_{20,A} \approx \left( \frac{30 \cdot T_{20} + 40 \cdot T_{10} + 30 \cdot T_{2k}}{90} - 0.06 \right)
\]

and

\[
EDT_{10,A} \approx \left( \frac{50 \cdot T_{2k} + 50 \cdot T_{4k}}{90} - 0.06 \right)
\]

where indices 500, 1k, 2k and 4k denote results at 500 Hz, 1 kHz, 2 kHz and 4 kHz octave bands respectively.

6. CONCLUSION

Late reverberation times do not correlate well to clarity or other monaural measures indeed. Yet they can still be used to estimate worst case limits of C50 or STI, which is exactly what a minimum specification should be about.

More general and emphasized use of early decay time is encouraged in general specifications, since it correlates well to clarity measures and perceived quality of a room as well.

Use of STI seems to be superfluous, as it is more complex to measure or calculate and is meant basically to qualify nonlinear and noisy transmission, none of which is a room acoustic feature. If it is important to express clarity in a simple way, EDT10 is more practical than STI or even C50.

Specifying spectral tolerances cannot be avoided if the source signal level is high in low frequency bands (e.g. music or amplification is active).

It seems adequate to specify quantity, uniform distribution and visibility of absorption to specify sound level reducing room acoustics measures.

Rooms shall be specified in their furnished, unoccupied condition, with special consideration if non-architectural objects affect initial room acoustic conditions when the room is inhabited.

7. ACKNOWLEDGEMENTS

The author would like to thank colleagues Éva Arató, Gergely Arató and Attila B. Nagy to provide additional room impulse response measurement data for this study.

The author would like to thank the Section for Acoustics of the Hungarian Engineering Chamber to support the preparation of a room acoustic design guideline, during which numerous observations were revealed.

8. REFERENCES


[7] C. Campbell, C. Svensson, E. Nilsson, „The same reverberation time in two identical rooms does not necessarily mean the same level of speech clarity and sound levels when we look at impact of different ceiling and wall absorbers”, Inter-noise 2014
